

53  
AI

1. A method of expanding an effective active vocabulary of a speech recognition system, the method comprising:

- using a speech recognizer to perform speech recognition on a user utterance to produce one or more recognition candidates, the speech recognition comprising comparing digital values representative of the user utterance to a set of acoustic models representative of an active vocabulary of the system, the set of acoustic models including models of words and models of word fragments,
- receiving the recognition candidates from the speech recognizer, and
- when a received recognition candidate includes a word fragment:
  - determining whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system;
  - if the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, modifying the recognition candidate to substitute the proposed word for the word fragment and the one or more adjacent word fragments or words used to form the proposed word; and
  - if the word fragment may not be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, discarding the recognition candidate.

1           2.    The method of claim 1, wherein the expanded  
2 effective vocabulary comprises words from the backup  
3 dictionary that are formed from a combination of words and  
4 word fragments or word fragments and word fragments from an  
5 active vocabulary that includes words and word fragments,  
6 and words from the active vocabulary.

1           3.    The method of claim 1, wherein word fragments  
2 comprise suffixes, prefixes, and roots that are not words.

1           4.    The method of claim 3, wherein:  
2           one or more spelling rules are associated with each  
3 prefix and each suffix,  
4           determining whether the word fragment may be  
5 combined with one or more adjacent word fragments or words  
6 to form a proposed word comprises using a prefix or suffix  
7 as the particular word fragment and using an associated  
8 spelling rule in forming the proposed word, and  
9           as a result of using the associated spelling rule, a  
10 spelling of the proposed word differs from a spelling that  
11 would result from merely concatenating the particular word  
12 fragment with the one or more adjacent word fragments or  
13 words.

1           5.    The method of claim 4, wherein determining  
2 whether the word fragment may be combined with one or more  
3 adjacent word fragments or words to form a proposed word  
4 comprises:  
5           retrieving from the received recognition candidate a  
6 sequence that includes the particular word fragment and  
7 adjacent word fragments or words; and  
8           determining if the sequence is a valid sequence.

1           6. The method of claim 5, wherein a valid sequence  
2 includes only one or more allowed adjacent combinations of  
3 word fragments and words.

1           7. The method of claim 6, wherein allowed adjacent  
2 combinations comprise one or more prefixes, followed by a  
3 root or a word, followed by one or more suffixes; a root or  
4 a word followed by one or more suffixes; and one or more  
5 prefixes followed by a root or a word.

1           8. The method of claim 6, wherein allowed adjacent  
2 combinations comprise one or more prefixes, followed by one  
3 or more roots or words, followed by one or more suffixes;  
4 one or more roots or words followed by one or more suffixes;  
5 and one or more prefixes followed by one or more roots or  
6 words.

1           9. The method of claim 4, further comprising  
2 combining the particular word fragment with the one or more  
3 adjacent word fragments or words to form a second proposed  
4 word that differs from the first proposed word by using a  
5 second associated spelling rule in forming the proposed  
6 word.

1 10. The method of claim 1, wherein:  
2 one or more spelling rules are associated with a  
3 particular word fragment,  
4 combining the particular word fragment with one or  
5 more adjacent word fragments or words to form a proposed  
6 word comprises using an associated spelling rule in forming  
7 the proposed word, and  
8 as a result of using the associated spelling rule, a  
9 spelling of the proposed word differs from a spelling that  
10 would result from merely concatenating the particular word  
11 fragment with the one or more adjacent word fragments or  
12 words.

1 11. The method of claim 1, wherein determining  
2 whether the word fragment may be combined with one or more  
3 adjacent word fragments or words to form a proposed word  
4 included in a backup dictionary of the speech recognition  
5 system comprises searching the backup dictionary for the  
6 proposed word.



664060 040000

1 15. The method of claim 14, wherein producing the  
2 score associated with the modified recognition candidate  
3 comprises combining the acoustic component of the score for  
4 the received recognition candidate with the language model  
5 score generated for the modified recognition candidate.

1 16. The method of claim 14, wherein rescoreing the  
2 modified recognition candidate comprises generating an  
3 acoustic model score for the modified recognition candidate  
4 and producing the score associated with the modified  
5 recognition candidate comprises combining the modified  
6 score generated for the modified recognition candidate with  
7 the language model score generated for the modified  
8 recognition candidate.

1 17. The method of claim 13, wherein:  
2 the score associated with the received recognition  
3 candidate includes an acoustic component and a language  
4 model component; and  
5 rescoreing the modified recognition candidate  
6 comprises generating an acoustic score for the modified  
7 recognition candidate.

1 18. The method of claim 17, wherein producing the  
2 score associated with the modified recognition candidate  
3 comprises combining the language model component of the  
4 score for the received recognition candidate with the  
5 acoustic score generated for the modified recognition  
6 candidate.



1           23. The method of claim 22, further comprising  
2 generating acoustic models for the roots using portions of  
3 acoustic models of the words of the backup dictionary that  
4 are not included in acoustic models of the affixes.

1           24. The method of claim 22, further comprising  
2 adding affixes and roots to the active vocabulary as word  
3 fragments.

1           25. The method of claim 24, further comprising  
2 storing a set of spelling rules in association with an affix  
3 in the active vocabulary.

1           26. The method of claim 24, further comprising  
2 creating a language model associated with the active  
3 vocabulary.

1           27. The method of claim 26, wherein creating the  
2 language model comprises:  
3           retrieving a training collection of text, the  
4 training collection of text comprising words from the backup  
5 dictionary and words from the active vocabulary;  
6           modifying the training collection of text by  
7 replacing any splittable backup dictionary words with their  
8 corresponding words and word fragments; and  
9           generating language model scores for words and word  
10 fragments of the active vocabulary using the modified  
11 collection of text.

1           28. The method of claim 26, wherein creating the  
2 language model comprises creating an N-gram language model.





1           33. The method of claim 1, wherein determining  
2 whether the word fragment may be combined with one or more  
3 adjacent word fragments or words to form a proposed word  
4 included in the backup dictionary of the speech recognition  
5 system comprises searching the backup dictionary for the  
6 proposed word based on a pronunciation of the proposed word.

33  
34. A method of recognizing speech, the method  
comprising:  
using a speech recognizer to perform speech  
recognition on a user utterance to produce a set of one or  
more recognition candidates, the speech recognition  
comprising comparing digital values representative of the  
user utterance to a set of acoustic models representative of  
an active vocabulary of the system, the set of acoustic  
models including models of words, models of roots that are  
not words, and models of affixes that are not words, the  
affixes including prefixes and suffixes,  
receiving the recognition candidates from the speech  
recognizer, and  
when a received recognition candidate includes an  
affix:  
combining the affix with one or more adjacent  
words, roots, or other affixes to form a new word; and  
modifying the recognition candidate to  
substitute the new word for the affix and the one or more  
adjacent words, roots, or other affixes used to form the new  
word.

1 35. A method of generating an acoustic model of a  
2 word fragment, the method comprising:  
3 comparing a word of an active vocabulary to a  
4 similar word of a backup dictionary to identify a word  
5 fragment that may be used to convert the word of the active  
6 vocabulary to the word of the backup dictionary, and  
7 generating the acoustic model of the word fragment  
8 using a portion of an acoustic model of the word of the  
9 backup dictionary that is not included in an acoustic model  
10 of the word of the active vocabulary.

1 36. The method of claim 35, wherein comparing a  
2 word of the active vocabulary to a similar word of a backup  
3 dictionary comprises comparing spellings of the two words.

1 37. A method of generating acoustic models of word  
2 fragments, the method comprising:  
3 comparing words of an active vocabulary to similar  
4 words of a backup dictionary to identify spelling rules that  
5 may be used to convert the words of the active vocabulary to  
6 words of the backup dictionary; and  
7 employing the spelling rules in identifying word  
8 fragments.

1 38. The method of claim 37, wherein employing the  
2 spelling rules in identifying word fragments comprises:  
3 grouping spelling rules together to form possible  
4 affixes, the affixes including prefixes and suffixes; and  
5 analyzing words of the backup dictionary using the  
6 affixes to identify roots that may be combined with the  
7 affixes to produce words of the backup dictionary.

1 39. The method of claim 38, further comprising  
2 generating acoustic models for the roots using portions of  
3 acoustic models of the words of the backup dictionary that  
4 are not included in acoustic models of the affixes.

1 40. The method of claim 38, further comprising  
2 adding affixes and roots to the active vocabulary as word  
3 fragments.

1 41. The method of claim 40, further comprising  
2 storing a set of spelling rules in association with an affix  
3 in the active vocabulary.

1 42. The method of claim 40, further comprising  
2 creating a language model associated with the active  
3 vocabulary.

1 43. The method of claim 42, wherein creating the  
2 language model comprises:  
3 retrieving a training collection of text, the  
4 training collection of text comprising words from the backup  
5 dictionary and words from the active vocabulary;  
6 modifying the training collection of text by  
7 replacing any splittable backup dictionary words with their  
8 corresponding words and word fragments; and  
9 generating language model scores for words and word  
10 fragments of the active vocabulary using the modified  
11 collection of text.

1 44. The method of claim 42, wherein creating the  
2 language model comprises creating an N-gram language model.

1           45. The method of claim 44, wherein creating the N-  
2 gram language model comprises:  
3           retrieving a training collection of text, the  
4 training collection of text comprising words from the backup  
5 dictionary and words from the active vocabulary;  
6           determining a frequency of each N-gram word sequence  
7 that appears in the training collection of text;  
8           modifying the N-gram word sequences by replacing any  
9 splittable backup dictionary words with their corresponding  
10 words and word fragments;  
11           based on the N-gram word sequence frequencies,  
12 determining a frequency of each modified N-gram sequence  
13 that includes words, word fragments, or words and word  
14 fragments; and  
15           based on the N-gram word and word fragment sequence  
16 frequencies, generating an N-gram language model for the  
17 words and word fragments of the active vocabulary.

46. A computer-implemented speech recognition system that uses an expanded effective active vocabulary, the system comprising:

- a storage device configured to store an active vocabulary that includes multiple entries corresponding to words, commands, and word fragments; and
- a processor configured to:
  - receive data representing a user utterance,
  - produce one or more recognition candidates, by comparing digital values representative of the user utterance to a set of acoustic models representative of the active vocabulary of the system, the set of acoustic models including models of words and models of word fragments,
  - when a produced recognition candidate includes a word fragment:
    - determine whether the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system;
    - if the word fragment may be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary, modify the recognition candidate to substitute the proposed word for the word fragment and the one or more adjacent word fragments or words used to form the proposed word; and
    - if the word fragment may not be combined with one or more adjacent word fragments or words to form a proposed word included in a backup dictionary of the speech recognition system, discard the recognition candidate.

1 47. The system of claim 46, wherein the expanded  
2 effective active vocabulary comprises words from the backup  
3 dictionary that are formed from a combination of words and  
4 word fragments or word fragments and word fragments from an  
5 active vocabulary that includes words and word fragments,  
6 and words from the active vocabulary.

1 48. The system of claim 46, wherein the processor  
2 determines whether the word fragment may be combined with  
3 one or more adjacent word fragments or words to form a  
4 proposed word included in the backup dictionary by searching  
5 the backup dictionary for the proposed word.

1 49. The system of claim 46, wherein the processor  
2 modifies the recognition candidate by:  
3 forming a prospective recognition candidate by  
4 modifying the recognition candidate to substitute the  
5 proposed word for the word fragment and the one or more  
6 adjacent word fragments or words used to form the proposed  
7 word; and

8 if the prospective recognition candidate includes an  
9 additional word fragment:

10 further processing the prospective recognition  
11 candidate to generate an additional word using the  
12 additional word fragment and one or more adjacent words or  
13 word fragments, and

14 forming a final recognition candidate by  
15 replacing the additional word fragment and the one or more  
16 adjacent words with the additional word.





51. Computer software, residing on a computer  
readable medium, for a speech recognition system that uses  
an expanded effective active vocabulary to recognize words,  
and commands, the computer software comprising instructions  
for causing a computer to perform the following operations:  
    receive data representing a user utterance,  
    produce one or more recognition candidates, by  
    comparing digital values representative of the user  
    utterance to a set of acoustic models representative of an  
active vocabulary of the system, the set of acoustic models  
including models of words and models of word fragments,  
    when a produced recognition candidate includes a  
word fragment:  
        determine whether the word fragment may be  
combined with one or more adjacent word fragments or words  
to form a proposed word included in a backup dictionary of  
the speech recognition system;  
        if the word fragment may be combined with one  
or more adjacent word fragments or words to form a proposed  
word included in a backup dictionary, modify the recognition  
candidate to substitute the proposed word for the word  
fragment and the one or more adjacent word fragments or  
words used to form the proposed word; and  
        if the word fragment may not be combined with  
one or more adjacent word fragments or words to form a  
proposed word included in a backup dictionary of the speech  
recognition system, discard the recognition candidate.

1           52. The computer software of claim 51, wherein the  
2 expanded effective active vocabulary comprises words from  
3 the backup dictionary that are formed from a combination of  
4 words and word fragments or word fragments and word  
5 fragments from an active vocabulary that includes words and  
6 word fragments, and words from the active vocabulary.

1           53. The computer software of claim 51, wherein  
2 determining whether the word fragment may be combined with  
3 one or more adjacent word fragments or words to form a  
4 proposed word included in the backup dictionary comprises  
5 searching the backup dictionary for the proposed word.

1           54. The computer software of claim 51, wherein  
2 modifying the recognition candidate comprises:  
3           forming a prospective recognition candidate by  
4 modifying the recognition candidate to substitute the  
5 proposed word for the word fragment and the one or more  
6 adjacent word fragments or words used to form the proposed  
7 word; and  
8           if the prospective recognition candidate includes an  
9 additional word fragment:  
10           further processing the prospective recognition  
11 candidate to generate an additional word using the  
12 additional word fragment and one or more adjacent words or  
13 word fragments, and  
14           forming a final recognition candidate by  
15 replacing the additional word fragment and the one or more  
16 adjacent words with the additional word.

1           55. The computer software of claim 51, wherein a  
2 score is associated with the received recognition candidate,  
3 the computer software comprising instructions for causing  
4 the computer to produce a score associated with the modified  
5 recognition candidate by rescoring the modified recognition  
6 candidate.

55. The computer software of claim 51, wherein a